

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 1999, 01:30:01 ; Search time 366.45 seconds  
(without alignments)  
1435.120 Million cell updates/sec

Title: US-09-205-015-2

Perfect score: 147

Sequence: 1 agataactggggccaaacatg.....ctctccctctctgtttatct 147

Scoring table: IDENTITY\_NUC

Searched: 808301 seqs, 178873984 residues

Database :

GenEmbl.\*  
1: gb\_ba1.\*  
2: gb\_ba2.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl1.\*  
9: gb\_pl2.\*  
10: gb\_pr1.\*  
11: gb\_pr2.\*  
12: gb\_pr3.\*  
13: gb\_ro.\*  
14: gb\_st.\*  
15: gb\_sy.\*  
16: gb\_un.\*  
17: gb\_v1.\*  
18: gb\_v2.\*  
19: gb\_v3.\*  
20: em\_fun.\*  
21: em\_hum1.\*  
22: em\_hum2.\*  
23: em\_in.\*  
24: em\_om.\*  
25: em\_ov.\*  
26: em\_pat.\*  
27: em\_ph.\*  
28: em\_pl.\*  
29: em\_ro.\*  
30: em\_sy.\*  
31: em\_un.\*  
32: em\_v1.\*  
33: em\_v2.\*  
34: em\_v3.\*  
35: em\_st.\*  
36: gb\_ba1.\*  
37: gb\_ba2.\*  
38: gb\_pl1.\*  
39: gb\_pl2.\*  
40: gb\_pr1.\*  
41: gb\_pr2.\*  
42: gb\_pr3.\*  
43: gb\_st.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
-----							

1	145.4	98.9	356	10	S49899	S49899 alpha-globi
2	145.4	98.9	19226	11	HSG64	284722 Human DNA s
3	145.4	98.9	356	40	S49899	S49899 alpha-globi
4	145.4	98.9	19226	41	HSG64	284722 Human DNA s
5	70.2	47.8	277	13	S78508	S78508 alpha globi
6	70.2	47.8	330	13	MMT08220	U08220 Mus musculu
7	31	21.1	9750	12	HSU1	U07378 Human colla
8	31	21.1	9750	42	HSU1	U07378 Human colla
9	30.6	20.8	1149	10	HSPKCG5	X62533 H.sapiens g
10	30.6	20.8	1149	40	HSPKCG5	X62533 H.sapiens g
11	30.4	20.7	40668	11	AC003112	AC003112 Human DNA
12	30.4	20.7	118313	11	AC003991	AC003991 Human DNA
13	30.4	20.7	229269	18	AC004953	AC004953 *** SEQUE
14	30.4	20.7	198776	18	AC005075	AC005075 *** SEQUE
15	30.4	20.7	40668	41	AC003112	AC003112 Human DNA
16	30.4	20.7	118313	41	AC003991	AC003991 Human DNA
17	30.2	20.5	5289	11	AB014533	AB014533 Homo sapi
18	30.2	20.5	115045	11	AB014533	AB014533 Homo sapi
19	30.2	20.5	5289	41	AB014533	AB014533 Homo sapi
20	30.2	20.5	115045	41	AB014533	AB014533 Homo sapi
21	30	20.4	59154	18	AC005440	AC005440 *** SEQUE
22	30	20.4	97904	18	AC005440	AC005440 *** SEQUE
23	29.6	20.1	4532	10	HUMPRPHX	L14565 Human perip
24	29.6	20.1	154405	12	AC004990	AC004990 Homo sapi
25	29.6	20.1	4532	40	HUMPRPHX	L14565 Human perip
26	29.6	20.1	154405	42	AC004990	AC004990 Homo sapi
27	29.4	20.0	36149	9	SPBC12C2	254140 S.pombe chr
28	29.4	20.0	36149	39	SPBC12C2	254140 S.pombe chr
29	29.2	19.9	89818	11	AC002126	AC002126 Homo sapi
30	29.2	19.9	50511	11	AC005214	AC005214 Homo sapi
31	29.2	19.9	34877	11	CH19R37740	AD000812 Homo sapi
32	29.2	19.9	71353	12	AF104455	AF104455 Homo sapi
33	29.2	19.9	1961	13	AB013455	AB013455 Rattus no
34	29.2	19.9	2440	13	RATNAPI2A	L13257 Rattus norv
35	29.2	19.9	2537	13	RATDROHOM	M90514 Rattus ratt
36	29.2	19.9	170891	18	AC002118	AC002118 *** SEQUE
37	29.2	19.9	40865	18	AC004195	AC004195 *** SEQUE
38	29.2	19.9	89818	41	AC002126	AC002126 Homo sapi
39	29.2	19.9	50511	41	AC005214	AC005214 Homo sapi
40	29.2	19.9	34877	41	CH19R37740	AD000812 Homo sapi
41	29.2	19.9	71353	42	AF104455	AF104455 Homo sapi
42	29	19.7	174311	3	AC004758	AC004758 Drosophill
43	29	19.7	5022	11	D83779	D83779 Human mRNA
44	29	19.7	97585	12	HS63G5	294160 Homo sapien
45	29	19.7	1315	13	MUSTCPMC	M22600 Mouse t com

## ALIGNMENTS

RESULT	1	S49899	356 bp	DNA	PRI	10-JUL-1992
LOCUS		S49899				
DEFINITION		alpha-globin gene cluster: [5' region, major regulatory element]				
ACCESSION		S49899				
NID		g233777				
KEYWORDS		human.				
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		1 (bases 1 to 356)				
AUTHORS		Higgin, D.R., Wood, W.G., Sharpe, J.A., Gourdon, G., Ayub, H. and				
TITLE		Characterization of the major regulatory element upstream of the				
JOURNAL		human alpha-globin gene cluster				
MEDLINE		Mol. Cell. Biol. 11 (9), 4679-4689 (1991)				
REMARK		91342671				
COMMENT		GenBank staff at the National Library of Medicine created this				
		entry [NCBI gbbsq 49899] from the original journal article.				
		This sequence comes from fig 5.				
		Region: alpha-globin gene cluster.				

```

FEATURES
  source
    Location/Qualifiers
      1..356
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
BASE COUNT      89 a   89 c   108 g   70 t
ORIGIN

Query Match      98.9%; Score 145.4; DB 10; Length 356;
Best Local Similarity 99.3%; Pred. No. 7.4e-38;
Matches 146; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 agataactggcccaacatgactcagctctctggaggcccaacaggactctgagtcac 60
    |||||||
Db 111 AGATAACTGGGCCAACCATGACTCAGCTCTCTGGAGGCCCAACAGGACTCTGAGTCATC 170

QY 61 ctgtgggggtggagggtgggcaagggaagggtgtaattgactctgattacaacctt 120
    |||||||
Db 171 CTGTGGGGGTGGAGTGGGACAAAGGAAGGGTGAATGCTACTGCTGATTACACCTCT 230

QY 121 ggtgctgcctccccctctctttatct 147
    |||||||
Db 231 GGTGCTGCCTCCCCCTCTCTTTATCT 257

RESULT 2
HSGG4      19226 bp      DNA      PRI      19-MAR-1997
LOCUS      Human DNA sequence from cosmid GG4 from a contig from the tip of
DEFINITION the short arm of chromosome 16, spanning 2Mb of 16p13.3.
ACCESSION  Z84722
NID        g1817579
KEYWORDS   16p13.3.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 19226)
AUTHORS   Flint, J. and Higgs, D.R.
TITLE     Direct Submission
JOURNAL   Submitted (28-JAN-1997) Sanger Centre, 970124/Hinxton,
COMMENT   Cambridgeshire, CB10 1RQ, UK. E-mail enquiries: humpub@sanger.ac.uk
          IMPORTANT: This sequence is not the entire insert of clone GG4.
          This clone was sequenced at the Institute of Molecular Medicine.
          The true left end of clone GG4 is at 1 in this sequence. The true
          right end of clone RA36 is at 456.
          The true left end of clone PX94 is at 19090.
          GG4 is from a 280kb clone contig extending from the telomere of
          16p.
          Higgs D.R., Flint J. unpublished. MRC Molecular Haematology Unit,
          Institute of Molecular Medicine, Oxford.
          GG4 came from the Los Alamos, flow sorted human Chromosome 16
          library.

FEATURES
  source
    Location/Qualifiers
      1..19226
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="16"
        /map="16p13.3"
        /clone="GG4"
        601..894
          /note="AluJb repeat: matches 301..5 of consensus"
          /note="AluJb repeat: matches 1..136 of consensus;
          incomplete repeat"
          1291..1601
            /note="AluSg repeat: matches 1..300 of consensus"
            /note="AluSg repeat: matches 132..288 of consensus;
            incomplete repeat"
            3487..3784
              /note="AluSg repeat: matches 1..299 of consensus"
              3809..4104
                /note="AluJo repeat: matches 1..302 of consensus"
                4823..5030
                  /note="AluSg repeat: matches 2..208 of consensus;
                  incomplete repeat"
                  5053..5092
                    /note="20 copies of 2 mer 85 % conserved"
                    5122..5280
                      /note="FAM repeat: matches 164..5 of consensus"
                      5759..5903
                        /note="L1MC2 repeat: matches 169..321 of consensus"
                        6130..6427
                          /note="AluSg repeat: matches 297..1 of consensus"
                          6428..6724
                            /note="AluJo repeat: matches 299..1 of consensus"
                            7144..7265
                              /note="L1MC2 repeat: matches 293..169 of consensus"
                              7433..7747
                                /note="AluSg repeat: matches 1..303 of consensus"
                                complement(join(7979..8064,10081..10234,14760..14834,
                                16585..16668))
                                /codon_start=1
                                /product="14 gene"
                                /db_xref="PID:g1817580"
                                /translation="MCGQKFLKIDNRFVGHPTLLQHALGQISKTDPSFKREAPMTI
                                LFNVFALRANADPSVINCILNLSRRIATVLOHEERRCQYLTRAKLILALQDEVSA
                                M
                                ADNEGQSPSPFHILPKCLARDLKEAYDR"
                                8094..8307
                                  /note="AluSg repeat: matches 212..1 of consensus;
                                  incomplete repeat"
                                  8578..8861
                                    /note="AluJb repeat: matches 3..297 of consensus"
                                    8907..9210
                                      /note="AluJb repeat: matches 300..1 of consensus"
                                      9807..9887
                                        /note="MIR repeat: matches 105..185 of consensus"
                                        12438..12731
                                          /note="AluJb repeat: matches 297..3 of consensus"
                                          12946..13243
                                            /note="AluSg repeat: matches 299..2 of consensus"
                                            14000..14139
                                              /note="AluJb repeat: matches 2..141 of consensus;
                                              incomplete repeat"
                                              14140..14453
                                                /note="AluSx repeat: matches 1..302 of consensus"
                                                14454..14596
                                                  /note="L1MB6 repeat: matches 760..907 of consensus"
                                                  16385..16584
                                                    /note="10 copies of 20 mer 91 % conserved"
                                                    18800..19089
                                                      /note="AluSx repeat: matches 1..293 of consensus"
                                                      BASE COUNT      5242 a   4881 c   4603 g   4500 t
                                                      ORIGIN

Query Match      98.9%; Score 145.4; DB 11; Length 19226;
Best Local Similarity 99.3%; Pred. No. 9.1e-38;
Matches 146; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 agataactggcccaacatgactcagctctctggaggcccaacaggactctgagtcac 60
    |||||||
Db 11063 AGATAACTGGGCCAACCATGACTCAGCTCTCTGGAGGCCCAACAGGACTCTGAGTCATC 11122

QY 61 ctgtgggggtggagggtgggcaagggaagggtgtaattgactctgattacaacctt 120
    |||||||
Db 11123 CTGTGGGGGTGGAGTGGGACAAAGGAAGGGTGAATGCTACTGCTGATTACACCTCT 11182

QY 121 ggtgctgcctccccctctctttatct 147
    |||||||
Db 11183 GGTGCTGCCTCCCCCTCTCTTTATCT 11209

RESULT 3

```

```

S49899          356 bp      DNA      PRI      10-JUL-1992
LOCUS          alpha-globin gene cluster: {5' region, major regulatory element}
DEFINITION     [human, Genomic, 356 nt].
ACCESSION      S49899
NID            9233777
KEYWORDS       human.
SOURCE         Homo sapiens
ORGANISM       Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
REFERENCE      1 (bases 1 to 356)
AUTHORS        Jarman,A.P., Wood,W.G., Sharpe,J.A., Gourdon,G., Ayyub,H. and
                Higgs,D.R.
TITLE          Characterization of the major regulatory element upstream of the
                human alpha-globin gene cluster
JOURNAL        Mol. Cell. Biol. 11 (9), 4679-4689 (1991)
MEDLINE        91342671
REMARK         GenBank staff at the National Library of Medicine created this
                entry [NCBI gibbsq 49899] from the original journal article.
COMMENT        This sequence comes from fig 5.
FEATURES       Region: alpha-globin gene cluster.
                Location/Qualifiers
                1..356
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                89 a      89 c      108 g      70 t
BASE COUNT     89 a      89 c      108 g      70 t
ORIGIN
Query Match      98.9%; Score 145.4; DB 40; Length 356;
Best Local Similarity 99.3%; Pred. No. 7.4e-38;
Matches 146; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 agataactggggcccaaccatgactcagtcgtctctggaggcccaacagcactcttgatcctc 60
Db 111 AGATAACTGGGCCCAACCATGACTCAGTCGTCCTGGAGGCCCAACGAGCTCTGAGTCATC
Qy 61 ctgtgggggtggaggtgggacaaaggggaaggggtgaatgctgctgattacaacctct 120
Db 171 CTGTGGGGGTGGAGGTGGGCAAGGGAAGGGGTGAATGCTGCTGATTACAACCTCT 230
Qy 121 ggtgctgctccctccctcttattct 147
Db 231 GGTCCTCCCTCCCTCCCTCTCTTTATCT 257

RESULT 4
HSGG4
LOCUS          HSGG4
DEFINITION     Human DNA sequence from cosmid GG4 PRI 19-MAR-1997
                the short arm of chromosome 16, spanning 2Mb of 16p13.3.
ACCESSION      284722
NID            91817579
KEYWORDS       16p13.3.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
                Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                1 (bases 1 to 19226)
AUTHORS        Flint,J. and Higgs,D.R.
TITLE          Direct Submission
JOURNAL        Submitted (28-JAN-1997) Sanger Centre, 970124/Hinxton,
                Cambridgeshire, CB10 1RQ, UK. E-mail enquires: humpub@sanger.ac.uk
COMMENT        IMPORTANT: This sequence is not the entire insert of clone GG4.
                This clone was sequenced at the Institute of Molecular Medicine.
                The true left end of clone GG4 is at 1 in this sequence. The true
                right end of clone R336 is at 456.
                The true left end of clone PX94 is at 19090.
                GG4 is from a 280kb clone contig extending from the telomere of
                16p.
                Higgs D.R., Flint J. unpublished. MRC Molecular Haematology Unit,
                Institute of Molecular Medicine, Oxford.

```

---

```

GG4 came from the Los Alamos, flow sorted human Chromosome 16
library.
FEATURES       source
                Location/Qualifiers
                1..19226
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="16"
                /map="16p13.3"
                /clone="GG4"
                601..894
                /note="AluJb repeat: matches 301..5 of consensus"
                1156..1290
                /note="AluSx repeat: matches 1..136 of consensus;
                incomplete repeat"
                1291..1601
                /note="AluSp repeat: matches 1..300 of consensus"
                1602..1758
                /note="AluSg repeat: matches 132..288 of consensus;
                incomplete repeat"
                3487..3784
                /note="AluSg repeat: matches 1..299 of consensus"
                3809..4104
                /note="AluJb repeat: matches 1..302 of consensus"
                4823..5030
                /note="AluSg repeat: matches 2..208 of consensus;
                incomplete repeat"
                5053..5092
                /note="20 copies of 2 mer 85 % conserved"
                5122..5280
                /note="FAM repeat: matches 164..5 of consensus"
                5759..5903
                /note="L1MC2 repeat: matches 169..321 of consensus"
                6130..6427
                /note="AluSg repeat: matches 297..1 of consensus"
                6428..6724
                /note="AluJb repeat: matches 299..1 of consensus"
                7144..7265
                /note="L1MC2 repeat: matches 293..169 of consensus"
                7453..7747
                /note="AluSp repeat: matches 1..303 of consensus"
                complement(join(7979..8064,10081..10234,14760..14834,
                16385..16668))
                /codon_start=1
                /product="14 gene"
                /db_xref="PID:e300374"
                /db_xref="PID:g1817580"
                /translation="MCGQKFKLDINRVFVGHPTLQHALGQISKTDPPKREAPTM
                LFNVFEALRANADPSVINCLHNSRIATVQHERRCQYLTREAKLILALQDEVSAM
                ADNEGQPSPFHHLPCKLRDLKEAYDR"
                8094..8307
                /note="AluSg repeat: matches 212..1 of consensus;
                incomplete repeat"
                8578..8861
                /note="AluJ repeat: matches 3..297 of consensus"
                8907..9210
                /note="AluJb repeat: matches 300..1 of consensus"
                9807..9887
                /note="MIR repeat: matches 105..185 of consensus"
                12438..12731
                /note="AluJb repeat: matches 297..3 of consensus"
                12946..13243
                /note="AluSp repeat: matches 299..2 of consensus"
                14000..14139
                /note="AluJb repeat: matches 2..141 of consensus;
                incomplete repeat"
                14140..14453
                /note="AluSx repeat: matches 1..302 of consensus"
                14454..14596
                /note="L1MB6 repeat: matches 760..907 of consensus"
                16385..16584
                /note="10 copies of 20 mer 91 % conserved"
                18800..19089
                /note="AluSx repeat: matches 1..293 of consensus"

```

BASE COUNT	5242 a	4881 c	4603 g	4500 t	
ORIGIN					
Query Match	98.9%;	Score 145.4;	DB 41;	Length 19226;	
Best Local Similarity	99.3%;	Pred. No. 9.1e-38;			
Matches 146;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	agataactgggccaacatgactcagtgctctctggaggccaacaggactcttgagtcac	60		
DB	11063	AGATAACTGGGCCAACCATGACTCAGTGCTTCTGGAGGCCAACAGACTGCTGAGTCATC	11122		
QY	61	ctgtgggggtgagtgaggacaagggaagggtgaatgtagtactgctgattacaacctct	120		
DB	11123	CTGTGGGGGTGAGGTGGGACAGGGGAAGGGTGAATGTTACTGCTGATTACAACCTCT	11182		
QY	121	ggtgctgctccctccctctgtttatct	147		
DB	11183	GGTGTGCTCCCTCCCTCTGTTTATCT	11209		
RESULT	5				
LOCUS	S78508	277 bp	DNA	ROD	26-SEP-1995
DEFINITION	alpha globin [regulatory element] [mice, Genomic, 277 nt].				
ACCESSION	S78508				
NID	9999308				
KEYWORDS	Mus sp.				
SOURCE	Mus sp.				
ORGANISM	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 277)				
AUTHORS	Gourdon,G., Sharpe,J.A., Higgs,D.R. and Wood,W.G.				
TITLE	The mouse alpha-globin locus regulatory element				
JOURNAL	Blood 86 (2), 766-775 (1995)				
MEDLINE	95329736				
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 168417] from the original journal article. This sequence comes from Fig. 2B.				
FEATURES	Location/Qualifiers				
source	1..277				
gene	/organism="Mus sp."				
	/db_xref="taxon:10095"				
	1..277				
	/partial				
	/gene="alpha globin"				
BASE COUNT	79 a	65 c	71 g	62 t	
ORIGIN					
Query Match	47.8%;	Score 70.2;	DB 13;	Length 277;	
Best Local Similarity	72.1%;	Pred. No. 3.7e-13;			
Matches 106;	Conservative	0;	Mismatches 38;	Indels 3;	Gaps 1;
QY	1	agataactgggccaacatgactcagtgctctctggaggccaacaggactcttgagtcac	60		
DB	111	AGATAACTAAGCCAAGCATGACTCAGAGTTCCTAGAGGCCACTAGGACTGCTGAGTAATA	170		
QY	61	ctgtgggggtgagtgaggacaagggaagggtgaatgtagtactgctgattacaacctct	120		
DB	171	CTTGGGGGTACAGAGTCAGAAA---GGAAGGACAAATGGTACCACCTGATTAGGACCTCT	227		
QY	121	ggtgctgctccctccctctgtttatct	147		
DB	228	GAGCGTGTTCCTCCATCCCTGTTATT	254		
RESULT	6				
LOCUS	MMU08220	330 bp	DNA	ROD	29-NOV-1995
DEFINITION	Mus musculus C57BL alpha-globin locus control region.				
ACCESSION	U08220				
NID	9507452				
KEYWORDS	house mouse.				
SOURCE	Mus musculus				
ORGANISM	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 330)				
AUTHORS	Kielman,M.F., Smits,R. and Bernini,L.F.				
TITLE	Localization and characterization of the mouse alpha-globin locus control region				
JOURNAL	Genomics 21 (2), 431-433 (1994)				
MEDLINE	94375071				
REFERENCE	2 (bases 1 to 330)				
AUTHORS	Kielman,M.F.				
TITLE	Direct Submission				
JOURNAL	Submitted (31-MAR-1994) Kielman M.F., State University Leiden, Human Genetics / Medical Genetic Center, Wassenaarseweg 72, Leiden, Zuid-Holland, The Netherlands, 2333 AL				
FEATURES	Location/Qualifiers				
source	1..330				
	/organism="Mus musculus"				
	/strain="C57BL"				
	/db_xref="taxon:10090"				
	/chromosome="11"				
	/map="Hba-locus"				
	1..330				
	/standard_name="alpha-globin locus control region"				
	/function="regulates expression of the mouse alpha-globin genes"				
	/label="LCR	94 a	80 c	79 g	77 t
BASE COUNT	94 a	80 c	79 g	77 t	
ORIGIN					
Query Match	47.8%;	Score 70.2;	DB 13;	Length 330;	
Best Local Similarity	72.1%;	Pred. No. 3.7e-13;			
Matches 106;	Conservative	0;	Mismatches 38;	Indels 3;	Gaps 1;
QY	1	agataactgggccaacatgactcagtgctctctggaggccaacaggactcttgagtcac	60		
DB	105	AGATAACTAAGCCAAGCATGACTCAGAGTTCCTAGAGGCCACTAGGACTGCTGAGTAATA	164		
QY	61	ctgtgggggtgagtgaggacaagggaagggtgaatgtagtactgctgattacaacctct	120		
DB	165	CTTGGGGGTACAGAGTCAGAAA---GGAAGGACAAATGGTACCACCTGATTAGGACCTCT	221		
QY	121	ggtgctgctccctccctctgtttatct	147		
DB	222	GAGCGTGTTCCTCCATCCCTGTTATT	248		
RESULT	7				
LOCUS	HSU1	9750 bp	mRNA	PRI	29-APR-1997
DEFINITION	Human collagen type XII alpha-1 precursor (COL12A1) mRNA, complete cds.				
ACCESSION	U73778				
NID	g1846004				
KEYWORDS	1 of 2				
SEGMENT	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 9750)				
AUTHORS	Gerecke,D.R., Olson,P.F., Koch,M., Knoll,J.H., Taylor,R., Hudson,D.R., Champlaud,M.F., Olsen,B.R. and Burgeson,R.E.				
TITLE	Complete primary structure of two splice variants of collagen XII, and assignment of alpha 1(XII) collagen (COL12A1), alpha 1(IX) collagen (COL9A1), and alpha 1(XIX) collagen (COL19A1) to human				

chromosome 6q12-q13  
Genomics 41 (2), 236-242 (1997)  
MEDLINE  
REFERENCE  
2 (bases 1 to 9750)  
Olson, P.  
Direct Submission  
Submitted (09-OCT-1996) Cutaneous Biology Research Center,  
Massachusetts General Hospital, Third floor, Bldg.149, Thirteenth  
St., Charlestown, MA 02129, USA  
Location/Qualifiers  
1. .9750  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/cell\_type="primary foreskin fibroblasts; primary foreskin  
keratinocytes; squamous cell carcinoma cells; endothelial  
(HUVEC) cell lines"  
1. .113  
/gene="COL12A1"  
114. .185  
/gene="COL12A1"  
114. .9305  
/note="member of the FACIT group of collagens"  
/codon\_start=1  
/product="collagen type XII alpha-1 precursor"  
/db\_xref="PID:g1846005"  
/translation="MRSRLPPLAALGAALLLSIAEAVDVPDPSDLNFKLIIDNTVHMS  
WAEVPGIVGIRITVDTPTDPTGKTEIQLSASTTETLLSELVETEVITVITVDEVEE  
SVPIGQITIGSTSPVEKPKGTEIKQCSVAWTDLVFLVDGSMVGRNNFYIL  
DFIALAASVADIGEKTGVVQVSSDTEFNLQYQDELLAIAIKPIKGGNTM  
TGADLYLVKFTESAGARVPKVAIITDGKSODEVEIPARELRNVGVFSLGI  
KAADAKELQASTPSLNHNVNANFADIVDIONELISQVSCVDSELGELVSGREV  
EPPSNLIAVESYKYLNNWNPSPVTKYKVALTNWAGSRQHALSPQPTTISVR  
DLSDRTFQISVAMKMTSSEPISEMEKTOPKQVQECNSGRDIRADIVFDGDIS  
IGIANVFKVRLVFLVKSFEISPNVQISLVQYSDPHTFTLKKPTVEDIIEAIN  
TFPFGVSTNGKMTVREKIFVPSKGSNVPKVMILITDGKSDAPDRDAIKLRN  
SDVEIFAVGVKDAVSELEAIAPAEHTVFTVEDEFAFORISFELTSCIRIEBEL  
AAIKKAVVPPKDLFSFSEVSYGFTKNWSPAGENVSYHITYKEAADDEVTWEPAS  
STVLSLSKPELYLVNVAEYEDGSIPLAGEETEEVKGAPRLKYTDDETDSFK  
ITWQACRLVRLIRIIPVAVGSESVTPPNORRTLENLPDKYVSVIPEVFS  
GPGPLIGNAATEVRGNRDLRVSDDTTSTMLKSWGAPKVKQILVITYTPVAGET  
QEVTVRGITTVIQLGKEGQYALSIVLSAGDALGFEGLTEERGSQDLVTK  
DITDTSIAGVTPAGMVGVRYSKSLYDQVDTGKNDPEAIHTMIENLQPTKYR  
ISVATYSSGGEPLTDATLSEKSLQVTDIIVLPYKMGKRLQSGSTASRFK  
RPHGRGQNAKVPPTVTSIVLKRQPTIYDITVLPYKMGKRLQSGSTASRFK  
SPRNLTSDPTMSFRVTPWEPAGEVGVKVTPTGDRRLGELVGVYDNTVLEE  
LRAGTYKVNVEGDEGESSPLVQEMTTLSDTVMPIILSSGMECLTAEADIVLV  
DGSWSIGANFTVRSPIRIVENFDIGPRVQIALAQISGDPTEWQLNAHDKSL  
LQAVNLPKGNLTGALNFIRQNFRTQAGMRPARKIGVLTIDGSDDDVEAPS  
KKLDEGVLEAIGKNADEVELMIATDPDHDYVADFSLRISVDLTINLCS  
VKGPDLEAPNLVISERTHSFVSWTPPSDSVDYKVEIYVPGSKRQEFYSRME  
TSTVLKDLPTETVYVNYVVEDEYSEPLKTEKTLFPVWSLNYDVTMTHYQW  
OPVGATYILSKPKVKDTEPRKEVRLGPTVNDMOLTDLPENTVAVTVQVADHL  
TSEPTVREVTLPRLPQDLKRDVTHSTWNVFVPEVPGVKRYIVRYKTPEDVKEV  
EVDRETSLSLKDLSQTLTVSVSAVHDEGESPPTAQETTRPVAPNLTKEVTS  
EGFRGESHGASVSLKRLTWGPRGSDKMETILNGENTLVFNLPNTIYVSTIA  
IYADESDHDLIGERTLPIITLQAPKSPRNQVYNATSLVNPNDPASGRVQYIR  
ITYQPTSGEHEQTITGRQNSVYLQKLPDPTTITVSSLYPDGEGVTRGRTK  
PLUNVRNLVNDPSTSTGNRWDAEGNPRQYKLFYAPAAAGPEELVPIPGNTNYAL  
RNLPQDTSYTVTVVPTVTEGGGTSDTGRMLRNLARNVQVNPTRNLGVSDGPNP  
GVLQYRVVYSPVDGTRPESIVVPGTRMVLRLPDLTXSNVLVLYSDGEGNPS  
PAGQRTLRSRPNRLRVEGTNLSLVANDHAGDQVQYRIIYSPWGPIDETVTPV  
GRNVIIQLOPDPYKTIIVIAIEDGGHLTGNGRTVGLLPSPQNIHSDWYTRF  
RWSDPSPSVPLGVKIVYKPVGSEMEAFVGMTSYTLNLPSTTINVDVAYDS  
GLVSPLETQGLTLYLNDLTKYIQWDTFCVRSRPHAAATSYRLKSLPADGTRGEI  
TVRGSETSGCTGLSPDIDYGVTVQTPNLEGGVSKVHTVKTPEATEPTPTPP  
PTIIPADVCKAKADIVFLTASWISGDDNKNVVKFINTVGGFDEISPAQIOW  
FVQYDSKVEKFLNTYNDKALGALQNIYRGNGRTGKALTFFKEKVLTWESGM  
KNPKVLVVDGSDQVKKALVLIQSGSFVYGVVADYDYNELANIASKPSERH  
FVDDPSFKEIEDNLITVCTEATISCPLEIYLDGYSFGPKMLEANLTKFNFSVQ  
GVLSDEGFSFSPSYAIRIQKNAFVNPQADLPHNGLPSPSYIILLFLPLPTSPFPAI  
WQITDRDYKQVGVIAIDPSSKTLTFFNNDTRGEVQTVTFTDEEYKTLFVGSFHKVHV

mat\_peptide 186. .9302  
/gene="COL12A1"  
/note="member of the FACIT group of collagens"  
/product="collagen type XII alpha-1"  
9306. .9750  
/gene="COL12A1"  
BASE COUNT 2885 a 2150 c 2372 g 2343 t  
ORIGIN

3'UTR

Query Match 21.1%; Score 31; DB 12; Length 9750;  
Best Local Similarity 56.1%; Pred. No. 3.4;  
Matches 78; Conservative 0; Mismatches 60; Indels 1; Gaps 1;

QY 5 aactggggcccaaccatgactcagtcgtcttctggaggcccaaggactctctgagtcactcgt 64  
Db 8987 ACCTGGGCTGGGGGGGCCAGGCTTCCGGGGCACACACAGGATGACGAGACCCCTGG 9046

QY 65 ggggggtgagtgaggacaaagggtgaatggact-gctgattacaactcgtgt 123  
Db 9047 GGAACGAGGTTTGGCAGGAGAGAAAGGTAAGAGGGTACTGGATCTCTCAGGACCTCGGG 9106

QY 124 gctgctccccctcctcgtt 142  
Db 9107 GCTGCTGGGCCCCAGCT 9125

RESULT 8  
HSU1  
LOCUS  
DEFINITION Human collagen type XII alpha-1 precursor (COL12A1) mRNA, complete cds.  
ACCESSION U73778  
NID g1846004  
KEYWORDS i of 2  
SEGMENT human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 9750)  
AUTHORS Gerecke, D.R., Olson, P.F., Koch, M., Knoll, J.H., Taylor, R., Hudson, D.L., Champilaud, M.F., Olsen, B.R. and Burgeson, R.E.  
TITLE Complete primary structure of two splice variants of collagen XII, and assignment of alpha 1(XII) collagen (COL12A1), alpha 1(X) collagen (COL9A1), and alpha 1(XIX) collagen (COL19A1) to human chromosomes 6q12-q13  
JOURNAL Genomics 41 (2), 236-242 (1997)  
MEDLINE 97288521  
REFERENCE 2 (bases 1 to 9750)  
AUTHORS Olson, P.  
TITLE Direct Submission  
JOURNAL Submitted (09-OCT-1996) Cutaneous Biology Research Center, Massachusetts General Hospital, Third floor, Bldg.149, Thirteenth St., Charlestown, MA 02129, USA  
Location/Qualifiers  
1. .9750  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/cell\_type="primary foreskin fibroblasts; primary foreskin keratinocytes; squamous cell carcinoma cells; endothelial (HUVEC) cell lines"  
1. .113

5'UTR

sig\_peptide

CDS

```
/gene="COL12A1"
114..185
/gene="COL12A1"
114..9305
/notes="member of the FACIT group of collagens"
/codon_start=1
/product="collagen type XII alpha-1 precursor"
/db_xref="PDB:1H846005"
/translat="MRSRLPPALAAALGAALLSSIEAEVDPSPDLNFKIIDNTVHMS
WAEVDPVIGYRITVDTPTKEFTLLSELVPEVETVITISYDEVEE
SVFVIGLTIQTSSKPKVEKPKGKTEIOKQSAWMDLVFLVDSWGRNNFKYIL
DETAALYSAEDIGEEKTRGVQVYSSDTRFENLNOYORDELLAAKKIPYKGGNTM
TGDADILVANTTESAGARVPKVAIILIDGKSDVEIPEARELRNVGVESVGI
KAADAKELQIASPISLNHFNVANFADAIQNELISQVCSGVDQLGELVSGEYV
EPPNLIAMQSVSKYKVLNPNPSPVPTGYKILTPMTAGSRHSLVSGPQTLLSVR
DLSADTEYQISVAMKGMTSEPTISIMEKTPQMKVQECRSGVDIKADIVFLVDSYS
IGIANFVQVLEFLVKFSEIPNRVQISLVOYSRDPHTEFTLKKTVEVDIEAIN
TFPYRGSGTNGKAMTVREKIPVPSGRSNPKVMILITDGKSDAFRPAIKLRN
SDVEIFAVGVDAVRSLEALASPPATHVFTVEDEDAFORISFELTOSICLRIGEL
AAIKKAYVPPKDLSEFSEVTSYGFKNWSPAGENVSIIHTYKEAAGDEVTVVERAS
STSVLSLKPETILLYNVAAEYEDGSEIPLAETTEEVKGAPRNKLVTDITDFSK
ITWQAPEGRVLRIRIIVPVAGGESREVTPPNQRRTELENLPDKTVEVSYPEFS
GPGPLTGNATERNRPNRDLRVSDPTTSMKLSWAGPKVKQVILVTVYVAGET
QEVTVRGTNTVILQIGKEGTVALSVTALYASGADALFEGCTILLESGSPDLVTK
DITDTSIGATWSAPMVRGVRSWSLSLDVDTGKELNPEDAIHTMIENLQPELVYK
ISVFATYSSEGEPELGDATLSELSQDSKTLKYDEETENTMRTWTKPAPKVNRYVY
RPHGGRQVAKVPPPTVTLKRLQPTTIDVLYPKMGEGKLRQKSGVTSARFK
SPRNLTQSDTSMFSEVTVWEPAGEVKGVTFTPTGDDRRGLGELVGVONTVLEE
LRAGTIVYVFGMFDGESPLVQEMTILSDTVMPIILSGMECLITRAEDIVLLY
DGSISGRANFTRVRSIRIVEFDIGPKRVOIALAOYSGDPTWQLNHARDKSL
LOANVILYGGNTLTGMALNFIRQONFRTOAGMRPARKIGVILITDGKSODDEKPS
KLDKDEGVLEFALIGIKNADEVELKMIATDDDDHDYNADEFESLSRVLDDLTNLNS
VKGPDLEAPNSLVNYSRHSFRVSWTPSPDSVDKRYKEIIPVSGKRGREFVYSME
TSTVLDKLEPTEVVNYSVVEDEYSEPLKGTETKLPVVPVSLNIDVGVPTMHVOM
QPVGATGYILSYKPVKDETEPREKVLGTVDNMQLTDLNPNTEYAVTVQVLEHD
TSEPVTREVTLPLRPQDLKRDVTHSTNMVFWEPVPGVKRYIVRYKTPEDVKEV
EVDRESSTSLKDLFSOTLVSVSAVDEGESPPVTAQETRPVAPTNLKITEVTS
EGPRGWDHGSADSVLRITWPGSGSDKMETILNGDENTLVFNENPNPTIYEVSTA
IYADESSDLDIGSERILPILITQAPRSGRNQVYNAISNLSLVKWDPASGRVQYR
ITYQPTSGEHEQPTIIGGRNSVVLQKLPKPTTITVSSLPDGEGRNTRGRTK
PLNTVRLNRYDPTSTLNVRMDHAGNPRQYKLFYAPAGGPEELVPIPGNTNYAIL
RNLOPTDYTVTVVTVTEGDDGRTSDTGRMLRGLARNVQVYVNTPNRLGVWRDPAP
GVLQYRVVYSPVDGTRPSESVIIVGPNRMVHLERLIPDTLYSNLVALLYSDGEGNPS
PAQGTPLRSPGRNLRVGETNLSLVADWDHAGPVOQYRIIISPTVGDPIDEYTVV
GRNVILQPLQPTPYKITIAYVEDGDGHLTGRTVGLLPPQNIHISDEYTFE
RVSDPSPQPLGVKITVYKPGSEMEAFVGMTSTLHNLNPNSTYDYNVYQAIDS
GLSVPLTDQGTLLYNVLDLYQIGWDTFCVKWSPHRAATSYRLKLSPADTGRQEI
TVRGETSHCTGLSPDIDYGVTVVTPNLEGFSVKYKHTVYKTEAPTEPTPP
PTIIPARDVCKGAKADIVFLTDASIGDDFNKVKVIFNTVGFDEISPAQIQVS
FVQYSDEVKSEKLTNYDKALALGALQNIYRGNTRTKALTFIKELVLTWESGMR
KNPVLYVVDGRSQDSQKKAALVIOGSGFSVGVVGVADVNYELANIASKPSERV
FIVDFSEKEDENLITVCEATISSCPLIYLGYSFVGFKMLEAVNLEKFNASVQ
GVLSSEGSFYSAYRIQCNFVNOPTADLHPNGLPSPSYIILFLRLPETPDPFAL
WQITDRDKVQGVTVADSESKTLSEFNKDRTRGEVTVTFDEEVKTLFYGFSFKHVIV
VTSKSVKIYIDCEYIEEDIKKEAGNTIDGVEIILKLLKGERKSAFQIOSDFIVCS
VNTSRDCDPLSPRDEGCKPAFNSCTQDSVGPVGPVGPAGGAGKAGPGRGERS
GAIGPGRGDIGPVGOPPCPGPGLSLTDPGQGRMGKGDAGEPLGCPRTGTPGL
PGPFPMGPDGRTGTGDSANGPRGPRGSPGSPGVTGPGKPKGPDGHPGPGS
GLUGKGGORGDIASQNMHRAVQCEQLISQNMHFNQMLNQLPNDYQSSRNQPGPP
GPPGPGSAGAGVAGPGRGPFPGTQMGPPGGRGLPGEKGERGGRGPRGLGPP
PGPQGESRTGPGSGTSGRPPGPRGPNSGIQGPPGPGYCDSSQSCASIPYNGQSY
GSG"
186..9302
/gene="COL12A1"
/notes="member of the FACIT group of collagens"
/product="collagen type XII alpha-1"
9306..9750
/gene="COL12A1"
2885 a 2150 c 2372 g 2343 t
mat_peptide
3' UTR
BASE COUNT 2885 a 2150 c 2372 g 2343 t
ORIGIN
```

```
Query Match 21.1%; Score 31; DB 42; Length 9750;
Best Local Similarity 56.1%; Pred. No. 3.4;
Matches 78; Conservative 0; Mismatches 60; Indels 1; Gaps 1;
QY 5 aactgggccaaccatgactcagtcgtcttctggagggccaacaggacttctgagtcatactgt 64
Db 8987 ACCTGGGCTGGGGGGCGCCAGGCTTCCCGGCACACACAGGAGTCCAGGACCCCTGG 9046
QY 65 ggggtggaggtgggacaggaagggtgaatggtact-gctgattacaacacctgtgt 123
Db 9047 GGAACAGAGTTTGCAGGAGAGAAAGGTGAAAGGGGTACTGTAATCTTCAGGACCTCGGG 9106
QY 124 gctgctccccctctctgt 142
Db 9107 GCTGCTGGGCCCCCAGGT 9125
RESULT 9
HSPKCG5 HSPKCG5 1149 bp DNA PRI 08-AUG-1995
LOCUS H sapiens gene for protein kinase C gamma, upstream region.
DEFINITION X62533
ACCESSION X62533
NID 935498
KEYWORDS protein kinase C; serine/threonine protein kinase.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1149)
AUTHORS Mahajna,J., King,P., Parker,P. and Haley,J.
TITLE Autoregulation of cloned human protein kinase C beta and gamma gene
promoters in U937 cells
JOURNAL DNA Cell Biol. 14 (3), 213-222 (1995)
MEDLINE 95186059
REFERENCE 2 (bases 1 to 1149)
AUTHORS Haley,J.D.
TITLE Direct Submission
JOURNAL Submitted (12-JUL-1990) J.D. Haley, ONCOGENE SCIENCE INC., 350
COMMUNITY DRIVE, MANHASSET, NY 11030, USA
FEATURES
source
1..1149
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="leukocyte"
/clone_lib="lambda gt10"
/chromosome="16"
1147..1149
/gene="PKC-G"
1147..1149
/gene="PKC-G"
/codon_start=1
/product="protein kinase C gamma"
BASE COUNT 296 a 331 c 363 g 159 t
ORIGIN
Query Match 20.8%; Score 30.6; DB 10; Length 1149;
Best Local Similarity 58.5%; Pred. No. 4.1;
Matches 72; Conservative 0; Mismatches 49; Indels 2; Gaps 1;
QY 20 gactcagtgcttctggagggccaacaggacttctgagtcatactctgtgggagtgagg 79
Db 585 GACACAGAGCCCCCAGGCGCCCCACAGGACACACAGGACCTAGTGGGGAGGAAACGCG 644
QY 80 acaaggggaagggtgaatggtactgctgatta--caacctgtggtgcctccccctc 137
Db 645 GCAGGATGACAGATTGTCAGGGTGTGGGGGGGAGCCAGGCTCAGAGGATGCCCTCCCTC 704
QY 138 ctg 140
Db 705 CAG 707
```



```

to AA009412 ze82h02.r1 Soares fetal heart NBHH19W Homo
sapiens cDNA clone 365523 5' (62..144); 100% identity."
11127..11331
/notes="DDS similarity to AAL36115 zk90b04.r1 Soares
pregnant uterus NBHPU Homo sapiens cDNA clone 490063 5'
(193..394); 94% identity.--DDS similarity to AA009412
ze82h02.r1 Soares fetal heart NBHH19W Homo sapiens cDNA
clone 365523 5' (145..346); 96% identity.--(1121..11331)
DDS similarity to AA009693 ze82h02.s1 Soares fetal heart
NBHH19W Homo sapiens cDNA clone 365523 3' (429..307); 95%
identity.--(1127..11246) DDS similarity to AA047548
zf1se02.r1 Soares fetal heart NBHH19W Homo sapiens cDNA
clone 377018 5' (222..341); 97% identity.--(11061..11331)
predicted exon, program: graill2exons_human_1.3, frame: 0,
quality: good, score: 73.000--(11061..11209) DDS
similarity to AA452628 zx33f04.r1 Soares total fetus
NB2HF8 9w Homo sapiens cDNA clone 788287 5' (167..313);
98% identity."
complement(11869..12161)
/rpt_family="Alu"
12502..12581
/notes="DDS similarity to AAL36115 zk90b04.r1 Soares
pregnant uterus NBHPU Homo sapiens cDNA clone 490063 5'
(395..477); 92% identity.--DDS similarity to AA009412
ze82h02.r1 Soares fetal heart NBHH19W Homo sapiens cDNA
clone 365523 5' (347..427); 94% identity.--DDS similarity
to AA009693 ze82h02.s1 Soares fetal heart NBHH19W Homo
sapiens cDNA clone 365523 3' (306..228); 99%
identity.--DDS similarity to AA450010 zx33f04.s1 Soares
total fetus NB2HF8 9w Homo sapiens cDNA clone 788287 3'
(309..229); 99% identity.--"
complement(12966..13240)
/rpt_family="Alu"
complement(13274..13531)
/rpt_family="Alu"
complement(13756..14029)
/notes="predicted exon, program: graill2exons_human_1.3,
frame: 2, quality: good, score: 61.000--Other overlapping
matches:--(13756..13826) DDS similarity to AA009412
ze82h02.r1 Soares fetal heart NBHH19W Homo sapiens cDNA
clone 365523 5' (428..496); 97% identity.--(13756..13984)
DDS similarity to AA009693 ze82h02.s1 Soares fetal heart
NBHH19W Homo sapiens cDNA clone 365523 3' (227..1); 98%
identity.--(13756..13984) DDS similarity to AA450010
zx33f04.s1 Soares total fetus NB2HF8 9w Homo sapiens cDNA
clone 788287 3' (228..1); 99% identity.--"
complement(14100..14672)
/rpt_family="Alu"
14885..15317
/notes="DDS similarity to AA406406 zvllie07.s1 Soares NBHMPu
S1 Homo sapiens cDNA clone 753348 3' (1..433); score: 858
identity: 431/433 (99%)--(14884..15237) DDS similarity
to W37175 zb21a02.r1 Soares fetal lung NBHL19W Homo
sapiens cDNA clone 302666 5' (1..355); 94%
identity.--(15227..14885) DDS similarity to AAL21532
zk89c11.s1 Soares pregnant uterus NBHPU Homo sapiens cDNA
clone 490004 3' (342..1); 99% identity.--(15227..14885)
DDS similarity to AAL27694 zk89c11.r1 Soares pregnant
uterus NBHPU Homo sapiens cDNA clone 490004 5' (126..467);
99% identity.--(15227..14897) DDS similarity to W46603
zc32h10.r1 Soares senescent fibroblasts NBHSF Homo sapiens
cDNA clone 324067 5' (328..1); 98%
identity.--(15227..15088) DDS similarity to W46604
zc32h10.s1 Soares senescent fibroblasts NBHSF Homo sapiens
cDNA clone 324067 3' (322..465); 96% identity."
15713..15760
/notes="DDS similarity to AAL21532 zk89c11.s1 Soares
pregnant uterus NBHPU Homo sapiens cDNA clone 490004 3'
(389..343); 99% identity.--DDS similarity to AAL27694
zk89c11.r1 Soares pregnant uterus NBHPU Homo sapiens cDNA
clone 490004 5' (77..125); 90% identity.--(15735..15713)
DDS similarity to W46603 zc32h10.r1 Soares senescent
fibroblasts NBHSF Homo sapiens cDNA clone 324067 5'

```

```

(351..329); 100% identity.--(15735..15713) DDS similarity
to W46604 zc32h10.s1 Soares senescent fibroblasts NBHSF
Homo sapiens cDNA clone 324067 3' (299..321); 100%
identity."
complement(15895..16082)
/notes="predicted exon, program: graill2exons_human_1.3,
frame: 0, quality: excellent, score: 81.000--Other
overlapping matches:--(15936..15895) DDS similarity to
AAL21532 zk89c11.s1 Soares pregnant uterus NBHPU Homo
sapiens cDNA clone 490004 3' (431..390); 100%
identity.--(15970..15895) DDS similarity to AAL27694
zk89c11.r1 Soares pregnant uterus NBHPU Homo sapiens cDNA
clone 490004 5' (1..76); 100% identity."
17926..18190
/rpt_family="Alu"
complement(18270..18438)
/notes="predicted exon, program: graill2exons_human_1.3,
frame: 0, quality: excellent, score:
100.000--(18438..18284) DDS similarity to W66776
mel7b11.r1 Soares mouse embryo NBME13.5 14.5 Mus musculus
cDNA clone 387741 5' similar to PIR:B38252 B38252
granulocyte colony-stimulating factor precursor
(157..1); 82% identity.--(18438..18406) DDS similarity to
AA049280 mj45d02.r1 Soares mouse embryo NBME13.5 14.5 Mus
musculus cDNA clone 479043 5' similar to SW:IL6B_MOUSE
Q00560 INTERLEUKIN-6 RECEPTOR BETA PRECURSOR (432..464);
88% identity."
complement(18540..18697)
/notes="predicted exon, program: graill2exons_human_1.3,
frame: 1, quality: excellent, score: 100.000--DDS
similarity to W66776 mel7b11.r1 Soares mouse embryo
NBME13.5 14.5 Mus musculus cDNA clone 387741 5' similar to
PIR:B38252 B38252 granulocyte colony-stimulating factor
receptor precursor (316..158); 92% identity.--DDS
similarity to AA049280 mj45d02.r1 Soares mouse embryo
NBME13.5 14.5 Mus musculus cDNA clone 479043 5' similar to
SW:IL6B_MOUSE Q00560 INTERLEUKIN-6 RECEPTOR BETA CHAIN
PRECURSOR (274..431); 93% identity.--"
complement(19080..19379)
/rpt_family="Alu"
complement(20080..20249)
/notes="predicted exon, program: graill2exons_human_1.3,
frame: 2, quality: excellent, score: 100.000--DDS
similarity to AA049280 mj45d02.r1 Soares mouse embryo
NBME13.5 14.5 Mus musculus cDNA clone 479043 5' similar to
SW:IL6B_MOUSE Q00560 INTERLEUKIN-6 RECEPTOR BETA CHAIN
PRECURSOR (104..273); 86% identity.--(20245..20080) DDS
similarity to W66776 mel7b11.r1 Soares mouse embryo
NBME13.5 14.5 Mus musculus cDNA clone 387741 5' similar to
PIR:B38252 B38252 granulocyte colony-stimulating factor
receptor precursor (482..317); 86% identity."
complement(20420..20549)
/notes="predicted exon, program: graill2exons_human_1.3,
frame: 0, quality: excellent, score:
100.000--(20522..20420) DDS similarity to AA049280
mj45d02.r1 Soares mouse embryo NBME13.5 14.5 Mus musculus
cDNA clone 479043 5' similar to SW:IL6B_MOUSE Q00560
INTERLEUKIN-6 RECEPTOR BETA CHAIN PRECURSOR -(1..103); 85%
identity.--(20549..20479) DDS similarity to AA042914
zk56f01.s1 Soares pregnant uterus NBHPU Homo sapiens cDNA
clone 486841 3' (338..410); 96% identity.--"
complement(20779..21043)
/rpt_family="Alu"
complement(21209..21494)
/notes="predicted exon, program: graill2exons_human_1.3,
frame: 0, quality: excellent, score: 94.000--Other
overlapping matches:--(21494..21213) DDS similarity to
AA042914 zk56f01.s1 Soares pregnant uterus NBHPU Homo
sapiens cDNA clone 486841 3' (55..337); 99%
identity.--(21362..21268) DDS similarity to W46603
zc32h10.r1 Soares senescent fibroblasts NBHSF Homo sapiens
cDNA clone 324067 5' (448..352); 97%
identity.--(21494..21268) DDS similarity to W46604

```



```

zc32h10.s1 Soares senescent fibroblasts NDHSF Homo sapiens
cDNA clone 59425 3' (73..298); 99% identity."
complement(21854..22765)
/rpt_family="Alu"
repeat_region
complement(22928..23336)
/rpt_family="Alu"
repeat_region
23983..24280
/rpt_family="Alu"
repeat_region
complement(24432..24658)
/rpt_family="Alu"
repeat_region
complement(32650..32696)
/rpt_family="Alu"
misc_feature
frame: 0, quality: good, score: 74.000"
/note="predicted exon, program: graill2exons_human_1.3,
34712..34777
/note="predicted exon, program: graill2exons_human_1.3,
frame: 1, quality: good, score: 69.000"
/note="predicted exon, program: graill2exons_human_1.3,
35520..35664
/note="predicted exon, program: graill2exons_human_1.3,
frame: 2, quality: excellent, score: 100.000--Other
overlapping matches--(35559..35664) DDS similarity to
R59871 yH07c12.r1 Homo sapiens cDNA clone 42571 5'
(1..107); 94% identity.--(35575..35664) DDS similarity to
R89669 ym97f05.r1 Homo sapiens cDNA clone 166881 5'
(1..90); 100% identity."
35752..35843
/note="predicted exon, program: graill2exons_human_1.3,
frame: 2, quality: excellent, score: 98.000--DDS
similarity to 59871 yH07c12.r1 Homo sapiens cDNA clone
42571 5' (108..199); 97% identity.--DDS similarity to
R89669 ym97f05.r1 Homo sapiens cDNA clone 166881 5'
(91..181); 97% identity.--DDS similarity to N99345
IMAGE:59425 Homo sapiens cDNA clone 59425 (17..111); 100%
identity."
complement(36643..36920)
/rpt_family="Alu"
repeat_region
complement(37271..37547)
/rpt_family="Alu"
misc_feature
37623..37775
/note="predicted exon, program: graill2exons_human_1.3,
frame: 2, quality: good, score: 65.000--DDS similarity to
59871 yH07c12.r1 Homo sapiens cDNA clone 42571 5'
(200..353); 99% identity.--DDS similarity to R9669
ym97f05.r1 Homo sapiens cDNA clone 166881 5' (182..337);
98% identity.--DDS similarity to N99345 IMAGE:59425 Homo
sapiens cDNA clone 59425 (112..264); 100%
identity.--(37624..37775) DDS similarity to AA424113
zv80e08.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA
clone 759974 5' (1..152); 99% identity."
complement(38029..38339)
/rpt_family="Alu"
repeat_region
38648..38750
/note="predicted exon, program: graill2exons_human_1.3,
frame: 1, quality: excellent, score: 84.000--DDS
similarity to N99345 IMAGE:59425 Homo sapiens cDNA clone
59425 (265..363); 96% identity.--DDS similarity to
AA424113 zv80e08.r1 Soares total fetus Nb2HF8 9w Homo
sapiens cDNA clone 759974 5' (153..254); 99%
identity.--DDS similarity to AA378559 EST91296 Synovial
sarcoma Homo sapiens cDNA 5' end (23..124); 100%
identity."
39803..39920
/note="DDS similarity to N99345 IMAGE:59425 Homo sapiens
cDNA clone 59425 (384..480); 99% identity.--DDS similarity
to AA378559 EST91296 Synovial sarcoma Homo sapiens cDNA 5'
end (125..242); 100% identity.--(39812..39941) predicted
exon, program: graill2exons_human_1.3, frame: 0, quality:
excellent, score: 94.000--(39870..39870) DDS similarity to
AA424113 zv80e08.r1 Soares total fetus Nb2HF8 9w Homo
sapiens cDNA clone 759974 5' (255..322); 100% identity."
40023..40140
/note="DDS similarity to N99345 IMAGE:59425 Homo sapiens
cDNA clone 59425 (481..598); 100% identity.--DDS
similarity to AA378559 EST91296 Synovial sarcoma Homo

```

```

sapiens cDNA 5' end (243..308); 94%
identity.--(40023..40102) DS similarity to AA424113
zv80e08.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA
clone 759974 5' (323..402); 100% identity.--(40081..40140)
predicted exon, program: graill2exons_human_1.3, frame: 0,
quality: marginal, score: 44.000"
complement(40353..40657)
/rpt_family="Alu"
repeat_region
BASE COUNT      8422 a 12040 c 11050 g 9156 t
ORIGIN

```

```

Query Match      20.7%; Score 30.4; DB 11; Length 40568;
Best Local Similarity 55.6%; Pred. No. 5.8;
Matches 58; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```

```

Qy 27 tgctctggaggccacaggaactcttgagtcacatcctgtggggtgagtggaagcaagg 86
      ||| ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 34851 TTCTCTGGATGACACAGGAGACTTTGTGGGGGTGTCCTGGGAGGTGGGGTCGGG 34910
      ||| ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 87 aaaggggtgaatggtactgtctgattacaacctcttggctgctgcct 130
      ||| ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 34911 AATGGGAGATCTCTCCAGTAGTGTGACGGGCTGGGGTTCCTCT 34954
      ||| ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 12
AC003991/c
LOCUS      AC003991      118313 bp      DNA      PRI      14-JAN-1998
DEFINITION Human BAC clone RG167B05 from 7q21, complete sequence.
ACCESSION      AC003991
NID      G2772535
KEYWORDS      HTG.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.

```

```

REFERENCE
1 (bases 1 to 118313)
Murray,J, Rohlfing,T and Antoniou,B.
The sequence of H. sapiens BAC clone RG167B05
Unpublished (1998)
REFERENCE
2 (bases 1 to 118313)
Waterston,R.
Direct Submission
Submitted (14-JAN-1998) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
SUBMITTED BY:
Genome Sequencing Center
Department of Genetics
Washington University
St. Louis MO 63108, USA
http://genome.wustl.edu/gsc
mailto:sapiens@watson.wustl.edu

```

```

COMMENT
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

```

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded or sequenced with an alternate  
chemistry; an attempt was made to resolve all sequencing problems,  
such as compressions and repeats; all regions were covered by  
sequence from more than one subclone; and the assembly was  
confirmed by restriction digest.

MAPPING INFORMATION:  
The sequence of this clone was established as part of a mapping and  
sequencing collaboration between the NHGRI Chromosome 7 Mapping  
Project (Eric D. Green, Director), John D. McPherson in the  
Department of Genetics (Washington University), and the Washington  
University Genome Sequencing Center. For additional information  
about the map position of this sequence, see  
http://www.nhgri.nih.gov/DIR/GTB/CHR7 or send

mailto:egreen@hgrl.nih.gov

#### SOURCE INFORMATION:

Clone RG167B05 is from the first release of the human BAC library C17B-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).

VECTOR: pBelOBAC11

Selection: chloramphenicol

#### NEIGHBORING SEQUENCE INFORMATION:

The actual start of this clone is at base position 1 of RG167B05. actual end of this clone is at base position 118313 of RG167B05.

This clone contains STS SWS2908 (NID:g1113638).

#### FEATURES

source

Location/Qualifiers

1..118313  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/clone="RG167B05"  
/clone\_lib="C17B-978SK-B"  
/map="7q21"

repeat\_region

/rpt\_family="MER1\_type"

repeat\_region

/rpt\_family="MER1\_type"

repeat\_region

/rpt\_family="MER1\_type"

repeat\_region

/rpt\_family="L2"

repeat\_region

/rpt\_family="L2"

repeat\_region

/rpt\_family="L2"

repeat\_region

/rpt\_family="Alu"

repeat\_region

/rpt\_family="Alu"

gene

complement(<5050..7954)

CDS

/gene="SRI"

/note="match to sorcin protein P30626 (PID:g267021);

H\_RG167B05.1"

/codon\_start=1

/product="calcium binding protein amplified in

mutlidrug-resistant cells"

/db\_xref="PID:g2772536"

/translation="MAYGHPGAGGGYPPGGYGAGGAPPGQTQDPLGYFAAVAG

ODGQIDADELQRLTQSGIAGGYR"

9625..10516

repeat\_region

/rpt\_family="L1"

repeat\_region

/rpt\_family="L1"

repeat\_region

/rpt\_family="MIR"

repeat\_region

/rpt\_family="MER1\_type"

repeat\_region

/rpt\_family="MER1\_type"

repeat\_region

/rpt\_family="MER1\_type"

repeat\_region

/rpt\_family="MIR"

repeat\_region

/rpt\_family="Alu"

repeat\_region

/rpt\_family="MER1\_type?"

repeat\_region

/rpt\_family="Alu"

repeat\_region  
16244..16326  
/rpt\_family="MER1\_type?"  
repeat\_region  
16327..16913  
/rpt\_family="L1"  
repeat\_region  
16912..17147  
/rpt\_family="L1"  
repeat\_region  
17199..17400  
/rpt\_family="MaLR"  
repeat\_region  
17401..17517  
/rpt\_family="MER1\_type?"  
repeat\_region  
18465..18536  
/rpt\_family="L1"  
repeat\_region  
18644..18689  
/rpt\_family="(CA)n"  
repeat\_region  
19101..19157  
/rpt\_family="MIR"  
repeat\_region  
19436..20346  
/rpt\_family="L1"  
repeat\_region  
20347..20650  
/rpt\_family="Alu"  
repeat\_region  
21158..21226  
/rpt\_family="L1"  
repeat\_region  
21425..21738  
/rpt\_family="Alu"  
repeat\_region  
21816..21903  
/rpt\_family="MER81"  
repeat\_region  
21910..22079  
/rpt\_family="MER1\_type"  
repeat\_region  
22715..23122  
/rpt\_family="L1"  
repeat\_region  
24950..25239  
/rpt\_family="Alu"  
repeat\_region  
25318..25618  
/rpt\_family="Alu"  
repeat\_region  
25673..25862  
/rpt\_family="L1"  
misc\_feature  
26008..26260  
/note="match to EST AA327226 (NID:g1979532)"  
repeat\_region  
26261..26341  
/rpt\_family="MaLR"  
misc\_feature  
26342..26414  
/note="match to EST AA327226 (NID:g1979532)"  
repeat\_region  
26564..26742  
/rpt\_family="L1"  
repeat\_region  
26853..27102  
/rpt\_family="L1"  
repeat\_region  
27108..27407  
/rpt\_family="Alu"  
repeat\_region  
28496..28701  
/rpt\_family="Alu"  
repeat\_region  
28961..29023  
/rpt\_family="(TGGG)n"  
repeat\_region  
29217..29523  
/rpt\_family="Alu"  
repeat\_region  
30548..30650  
/rpt\_family="MIR"  
repeat\_region  
32058..32121  
/rpt\_family="L1"  
repeat\_region  
32134..32276  
/rpt\_family="L1"  
repeat\_region  
33240..36502  
/rpt\_family="L1"  
repeat\_region  
36503..36633  
/rpt\_family="Alu"  
repeat\_region  
36634..37363  
/rpt\_family="L1"  
repeat\_region  
37394..37548  
/rpt\_family="MaLR"  
repeat\_region  
37657..38131  
/rpt\_family="L1"  
repeat\_region  
38133..38265  
/rpt\_family="Alu"  
repeat\_region  
38278..38556

```
repeat_region /rpt_family="Alu"
repeat_region 38570. .38737
repeat_region /rpt_family="Alu"
repeat_region 38746. .40503
repeat_region /rpt_family="L1"
repeat_region 40507. .40560
repeat_region /rpt_family="L1"
repeat_region 40573. .40874
repeat_region /rpt_family="Alu"
repeat_region 40948. .41237
repeat_region /rpt_family="Alu"
repeat_region 41457. .42699
repeat_region /rpt_family="L1"
repeat_region 43683. .44216
repeat_region /rpt_family="L1"
repeat_region 44330. .44975
repeat_region /rpt_family="L1"
repeat_region 45663. .46301
repeat_region /rpt_family="L2"
repeat_region 46475. .46665
repeat_region /rpt_family="Alu"
repeat_region 46691. .46867
repeat_region /rpt_family="Alu"
repeat_region 46875. .47123
repeat_region /rpt_family="L2"
repeat_region 47148. .47453
repeat_region /rpt_family="Alu"
repeat_region 47465. .47723
repeat_region /rpt_family="L2"
repeat_region 48947. .49001
repeat_region /rpt_family="MIR"
repeat_region 49438. .49710
repeat_region /rpt_family="Alu"
repeat_region 50605. .50902
repeat_region /rpt_family="Alu"
repeat_region 50923. .51302
repeat_region /rpt_family="L1"
repeat_region 51531. .51683
repeat_region /rpt_family="MER1_type"
repeat_region 51858. .52279
repeat_region /rpt_family="L1"
repeat_region 52333. .52634
repeat_region /rpt_family="Alu"
repeat_region 52761. .53284
repeat_region /rpt_family="MER4-group"
repeat_region 53422. .54408
repeat_region /rpt_family="GVA"
misc_feature 53986. .55215
repeat_region /note="CPG island (8CG-72.8, o/e=0.70, #CpGs=95)"
repeat_region 56698. .56786
repeat_region /rpt_family="Retroviral"
repeat_region 57713. .57791
repeat_region /rpt_family="MIR"
repeat_region 58071. .58124
repeat_region /rpt_family="L1"
repeat_region 58849. .59127
repeat_region /rpt_family="L1"
misc_feature 59146. .59176
misc_feature /note="match to EST AA493281 (NID:g2223122) nh55d02.s1"
misc_feature complement(59155. .59314)
repeat_region /note="match to EST AA493281 (NID:g2223122) nh55d02.s1"
repeat_region 59317. .59874
repeat_region /rpt_family="L1"
repeat_region 60088. .60425
repeat_region /rpt_family="Retroviral"
repeat_region 60906. .61152
repeat_region /rpt_family="L1"
repeat_region 61996. .62213
repeat_region /rpt_family="MIR"
repeat_region 62445. .63016
repeat_region /rpt_family="MER4-group"
repeat_region 63017. .63319
repeat_region /rpt_family="Alu"

repeat_region 63321. .63385
misc_feature /rpt_family="MER4-group"
misc_feature complement(64841. .64977)
misc_feature /note="match to EST R27235 (NID:g783370) yh53a03.r1"
misc_feature complement(64937. .65009)
misc_feature /note="match to EST R76842 (NID:g851474) y163c09.r1"
misc_feature complement(65019. .65370)
misc_feature /note="match to EST R76842 (NID:g851474) y163c09.r1"
repeat_region complement(65048. .65237)
repeat_region /note="match to EST R27235 (NID:g783370) yh53a03.r1"
repeat_region 65522. .65805
repeat_region /rpt_family="Alu"
repeat_region 68196. .68360
repeat_region /rpt_family="MER1_type"
repeat_region 70076. .70373
repeat_region /rpt_family="Alu"
misc_feature 70374. .70469
misc_feature /note="match to EST AA244027 (NID:g1874750) nc04a09.r1"
misc_feature complement(71252. .71561)
repeat_region /note="match to EST (NID:g966875)"
repeat_region 72533. .73353
repeat_region /rpt_family="L1"
repeat_region 73462. .73704
repeat_region /rpt_family="Alu"
repeat_region 73841. .73890
repeat_region /rpt_family="Retroviral"
repeat_region 73904. .74470
repeat_region /rpt_family="MER4-group"
repeat_region 74471. .74638
repeat_region /rpt_family="Alu"
repeat_region 74682. .74743
repeat_region /rpt_family="(CA)n"
repeat_region 75310. .76002
repeat_region /rpt_family="Retroviral"
repeat_region 76097. .76396
repeat_region /rpt_family="Alu"
repeat_region 76517. .76788
repeat_region /rpt_family="L1"
repeat_region 77027. .77111
repeat_region /rpt_family="L2"
repeat_region 78203. .78305
repeat_region /rpt_family="MIR"
repeat_region 78316. .78394
repeat_region /rpt_family="MER84"
repeat_region 78398. .78897
repeat_region /rpt_family="MER84"
repeat_region 79187. .79434
repeat_region /rpt_family="Alu"
repeat_region 80384. .81898
repeat_region /rpt_family="L1"
repeat_region 82043. .82107
repeat_region /rpt_family="MER2_type"
repeat_region 82259. .82394
repeat_region /rpt_family="MER2_type"
repeat_region 83316. .83472
repeat_region /rpt_family="MER1_type"
repeat_region 83526. .83705
repeat_region /rpt_family="MER1_type"
repeat_region 85044. .85127
repeat_region /rpt_family="L2"
repeat_region 85398. .85585
repeat_region /rpt_family="MER53"
repeat_region 85709. .86050
repeat_region /rpt_family="MER2_type"
repeat_region 86395. .86438
repeat_region /rpt_family="(CA)n"
repeat_region 88133. .88402
repeat_region /rpt_family="Alu"
repeat_region 88405. .88491
repeat_region /rpt_family="Alu"
repeat_region 88530. .88642
repeat_region /rpt_family="MER1_type"
repeat_region 89216. .89504
```

QY	87	aaagggggaatggtactgcctgattaca 114	Db	62935	ACAGGTCACAAAGACCTTGCTGATATAA 62908
repeat_region	/rpt_family="Alu"	89532..89803			
repeat_region	/rpt_family="Alu"	90600..90899			
repeat_region	/rpt_family="Alu"	91144..91440			
repeat_region	/rpt_family="Alu"	91948..92265			
repeat_region	/rpt_family="L1"	93862..94096			
repeat_region	/rpt_family="Alu"	95914..97065			
repeat_region	/rpt_family="MER2_type"	101098..101399			
repeat_region	/rpt_family="Alu"	101502..101762			
repeat_region	/rpt_family="L2"	103349..103798			
repeat_region	/rpt_family="L2"	103892..103943			
repeat_region	/rpt_family="(TGAG)n"	104105..104404			
repeat_region	/rpt_family="Alu"	104531..104893			
repeat_region	/rpt_family="L2"	104936..105236			
repeat_region	/rpt_family="Alu"	106604..106731			
repeat_region	/rpt_family="MER4-group"	106755..107066			
repeat_region	/rpt_family="Alu"	107090..107164			
repeat_region	/rpt_family="MER4-group"	107179..107530			
repeat_region	/rpt_family="MER4-group"	107625..108102			
repeat_region	/rpt_family="MaLR"	108218..108611			
repeat_region	/rpt_family="Retroviral"	109287..109762			
repeat_region	/rpt_family="MER1_type"	110537..110609			
repeat_region	/rpt_family="L2"	111360..111454			
repeat_region	/rpt_family="L2"	112493..112909			
repeat_region	/rpt_family="MaLR"	112938..112987			
repeat_region	/rpt_family="MIR"	114618..114797			
repeat_region	/rpt_family="MIR"	116497..116699			
repeat_region	/rpt_family="L1"	117098..117229			
repeat_region	/rpt_family="L1"	117857..118061			
repeat_region	/rpt_family="MER52"	118052..118167			
repeat_region	/rpt_family="MER52"	118202..118259			
repeat_region	/rpt_family="MER52"	118202..118259			
BASE COUNT	35024 a	22989 c	22511 g	37789 t	
ORIGIN					
Query Match	20.7%;	Score 30.4;	DB 11;	Length 118313;	
Best Local Similarity	59.1%;	Pred. No. 6.1;			
Matches	52;	Conservative	0;	Mismatches	36;
				Indels	0;
				Gaps	0;
QY	27	tgctctggagggcaacaggactcttgagtcctctgtgggtgtgggtgggagggg 86			
Db	62995	TGCATTCACAGGACGAGGCTGTTCTTAGTCACAGGATGAGTAGGTAAGCACAAGGT 62936			

```

* 108143 109610: contig of 1468 bp in length
* 109611 109625: gap of unknown length
* 111369 111383: contig of 1744 bp in length
* 111370 111384: gap of unknown length
* 111385 111399: contig of 1682 bp in length
* 113067 113081: gap of unknown length
* 113082 113096: contig of 1739 bp in length
* 114821 114835: gap of unknown length
* 114836 114850: contig of 1760 bp in length
* 116596 116610: gap of unknown length
* 116611 116625: contig of 1891 bp in length
* 118502 118516: gap of unknown length
* 118517 229269: contig of 110753 bp in length.
FEATURES
    source
        1. .229269
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="DJ1059M17"
BASE COUNT 60819 a 55407 c 54893 g 57758 t 392 others
ORIGIN

Query Match      20.7%; Score 30.4; DB 18; Length 229269;
Best Local Similarity 59.1%; Pred.No.6.3;
Matches 52; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 27 tgccttgaggccacagcagctctgagtcacctcctgtgggggtggaggtgggacaaagg 86
Db -22104 TGCATTCCAGGACAGCAGGTGTTCTTAGTCACAGGATGAGATAGGATAGACACAGGT 22045

Qy 87 aaaggggtgaatgtactgtgattaca 114
Db 22044 ACAGGTCAACAAGACCTTGCTGATAAAA 22017

RESULT 14
AC005075/c
LOCUS AC005075 198776 bp DNA HTG 12-JUN-1998
DEFINITION *** SEQUENCING IN PROGRESS *** Homo sapiens clone RG219E16; HTGS
phase 1, 3 unordered pieces.
ACCESSION AC005075
NID G3212916
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
Waterston,R.H.
1 (bases 1 to 198776)
The sequence of Homo sapiens clone
Unpublished
Waterston,R.H.
2 (bases 1 to 198776)
Direct Submission
Submitted (12-JUN-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
*** WARNING: Phase 1 High Throughput Genome Sequence ***
***
* This sequence is unfinished. It consists of 3 contigs for
* which the order is not known; their order in this record is
* arbitrary. In some cases, the exact lengths of the gaps
* between the contigs are also unknown; these gaps are presented
* as runs of N as a convenience only. When sequencing is complete,
* the sequence data presented in this record will be replaced
* by a single finished sequence with the same accession number.
* 1
* 1784: contig of 1784 bp in length
* 1785
* 1801: gap of unknown length
* 1802
* 18710: contig of 16909 bp in length
* 18711
* 18727: gap of unknown length
* 18728
* 198776: contig of 180049 bp in length.
    Location/Qualifiers
FEATURES
    source
        1. .198776
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="RG219E16"
BASE COUNT 60792 a 36795 c 37289 g 63866 t 34 others
ORIGIN

Query Match      20.7%; Score 30.4; DB 18; Length 198776;
Best Local Similarity 59.1%; Pred.No.6.3;
Matches 52; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 27 tgccttgaggccacagcagctctgagtcacctcctgtgggggtggaggtgggacaaagg 86
Db 183095 TGCATTCCAGGACAGCAGGTGTTCTTAGTCACAGGATGAGATAGGATAGACACAGGT 183036

Qy 87 aaaggggtgaatgtactgtgattaca 114
Db 183035 ACAGGTCAACAAGACCTTGCTGATAAAA 183008

RESULT 15
AC003112
LOCUS AC003112 40668 bp DNA PRI 21-NOV-1997
DEFINITION Human DNA from chromosome 19 specific cosmid R30292, genomic
sequence, complete sequence.
ACCESSION AC003112
NID G2636669
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
Lamerdin,J.E., McCreedy,P.M., Adamson,A.W., Burkhardt-Schultz,K.,
Gordon,L., Christensen,M., Kyle,A., Ramirez,M., Stillwagen,S.,
Garnes,J., Danganan,L., Bruce,R., Quan,G., Montgomery,M., Ow,D.,
Kobayashi,A., Olsen,A.O. and Carrano,A.V.
Sequence analysis of an ~1 Mb region containing the MEF2B gene in
19p12
Unpublished
2 (bases 1 to 40668)
Lamerdin,J.E.
Direct Submission
Submitted (21-NOV-1997) Human Genome Center, Lawrence Livermore
National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
Location/Qualifiers
    source
        1. .40668
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="R30292"
        /chromosome="19"
        /map="19p12 between UBA52 and D19S451"
        /cell_line="5HL2-B"
        /clone_lib="LL19NCO3 R chromosome 19 cosmid library"
        /note="LL19NCO3 cosmid library constructed at LLNL from
        flow-sorted chromosomes from hybrid 5HL2-B, which carries
        chromosome 19 as its only human chromosome."
        complement(6..62)
        /rpt_family="Alu"
        complement(571..640)
        /note="predicted exon, program: grill2exons_human_1.3,
        frame: 1, quality: good, score: 71.000"
        complement(1383..1444)
        /note="predicted exon, program: grill2exons_human_1.3,
        frame: 1, quality: good, score: 62.000"
        2257..8516
        /standard_name="endogenous retroviral sequence"
        /note="HERV9 retroviral sequence"
        complement(3269..3544)
        /rpt_family="Alu"
        5425..5722
        /rpt_family="Alu"
repeat_region
misc_feature
misc_feature
LTR
repeat_region
repeat_region

```

```

repeat_region complement(6486..6772)
/rpt_family="Alu"
repeat_region complement(7505..7783)
/rpt_family="Alu"
repeat_region 8951..9254
/rpt_family="Alu"
repeat_region 9358..9535
/rpt_family="LTR12"
misc_feature complement(9445..9505)
/note="predicted exon, program: grill2exons_human_1.3,
frame: 1, quality: good, score: 63.000"
misc_feature complement(9629..9672)
/note="predicted exon, program: grill2exons_human_1.3,
frame: 1, quality: excellent, score: 75.000"
repeat_region complement(9928..10123)
/rpt_family="Alu"
misc_feature 10409..10548
/note="DSS similarity to AA047548 zf15e02.rl Soares fetal
heart NBH19W Homo sapiens cDNA clone 377018 5' (1..138);
96% identity.--Other overlapping matches:--(10435..10548)
DSS similarity to AA136115 zk90b04.rl Soares pregnant
uterus NBHPU Homo sapiens cDNA clone 490063 5' (1..110);
93% identity.--(10466..10548) DDS similarity to AA452628
zx33f04.rl Soares total fetus NB2HF8 9w Homo sapiens cDNA
clone 788287 5' (1..82); 95% identity.--(10486..10548) DDS
similarity to AA009412 ze82h02.rl Soares fetal heart
NBH19W Homo sapiens cDNA clone 365523 5' (1..61); 97%
identity."
misc_feature 10643..10725
/note="predicted exon, program: grill2exons_human_1.3,
frame: 1, quality: excellent, score: 100.000--DSS
similarity to AA047548 zf15e02.rl Soares fetal heart
NBH19W Homo sapiens cDNA clone 377018 5' (139..221); 100%
identity.--DSS similarity to AA136115 zk90b04.rl Soares
pregnant uterus NBHPU Homo sapiens cDNA clone 490063 5'
(111..192); 99% identity.--DSS similarity to AA452628
zx33f04.rl Soares total fetus NB2HF8 9w Homo sapiens cDNA
clone 788287 5' (83..165); 100% identity.--DSS similarity
to AA009412 ze82h02.rl Soares fetal heart NBH19W Homo
sapiens cDNA clone 365523 5' (62..144); 100% identity."
misc_feature 11127..11331
/note="DSS similarity to AA136115 zk90b04.rl Soares
pregnant uterus NBHPU Homo sapiens cDNA clone 490063 5'
(193..394); 94% identity.--DSS similarity to AA009412
ze82h02.rl Soares fetal heart NBH19W Homo sapiens cDNA
clone 365523 5' (145..346); 96% identity.--(11121..11331)
DSS similarity to AA009693 ze82h02.s1 Soares fetal heart
NBH19W Homo sapiens cDNA clone 365523 3' (429..307); 95%
identity.--(11127..11246) DDS similarity to AA047548
zf15e02.rl Soares fetal heart NBH19W Homo sapiens cDNA
clone 377018 5' (222..341); 97% identity.--(11061..11331)
predicted exon, program: grill2exons_human_1.3, frame: 0,
quality: good, score: 73.000--(11061..11209) DDS
similarity to AA452628 zx33f04.rl Soares total fetus
NB2HF8 9w Homo sapiens cDNA clone 788287 5' (167..313);
98% identity."
repeat_region complement(11869..12161)
/rpt_family="Alu"
misc_feature 12502..12581
/note="DSS similarity to AA136115 zk90b04.rl Soares
pregnant uterus NBHPU Homo sapiens cDNA clone 490063 5'
(395..477); 92% identity.--DSS similarity to AA009412
ze82h02.rl Soares fetal heart NBH19W Homo sapiens cDNA
clone 365523 5' (347..427); 94% identity.--DSS similarity
to AA009693 ze82h02.s1 Soares fetal heart NBH19W Homo
sapiens cDNA clone 365523 3' (306..228); 99%
identity.--DSS similarity to AA450010 zx33f04.s1 Soares
total fetus NB2HF8 9w Homo sapiens cDNA clone 788287 3'
(309..229); 99% identity.--(13240)
complement(12966..13240)
/rpt_family="Alu"
repeat_region complement(13274..13531)
/rpt_family="Alu"
complement(13756..14029)
/note="predicted exon, program: grill2exons_human_1.3,
frame: 2, quality: good, score: 61.000--Other overlapping
matches:--(13756..13826) DDS similarity to AA009412
ze82h02.rl Soares fetal heart NBH19W Homo sapiens cDNA
clone 365523 5' (428..496); 97% identity.--(13756..13984)
DSS similarity to AA009693 ze82h02.s1 Soares fetal heart
NBH19W Homo sapiens cDNA clone 365523 3' (227..1); 98%
identity.--(13756..13984) DDS similarity to AA450010
zx33f04.s1 Soares total fetus NB2HF8 9w Homo sapiens cDNA
clone 788287 3' (228..1); 99% identity.--"
repeat_region complement(14100..14672)
/rpt_family="Alu"
misc_feature 14885..15317
/note="DSS similarity to AA046406 zvllc07.s1 Soares NBHMPU
S1 Homo sapiens cDNA clone 753348 3' (1..433); Score: 858
Identity: 431/433 (99%).--(14884..15237) DDS similarity
to W37175 zb21a02.rl Soares fetal lung NBH19W Homo
sapiens cDNA clone 302666 5' (1..355); 94%
identity.--(15227..14885) DDS similarity to AA121532
zk89c11.s1 Soares pregnant uterus NBHPU Homo sapiens cDNA
clone 490004 3' (342..1); 99% identity.--(15227..14885)
DSS similarity to AA127694 zk89c11.rl Soares pregnant
uterus NBHPU Homo sapiens cDNA clone 490004 5' (126..467);
99% identity.--(15227..14897) DDS similarity to W46603
zc32h10.rl Soares senescent fibroblasts NBHSF Homo sapiens
cDNA clone 324067 5' (328..1); 98%
identity.--(15227..15088) DDS similarity to W46604
zc32h10.s1 Soares senescent fibroblasts NBHSF Homo sapiens
cDNA clone 324067 3' (322..465); 96% identity."
misc_feature 15713..15760
/note="DSS similarity to AA121532 zk89c11.s1 Soares
pregnant uterus NBHPU Homo sapiens cDNA clone 490004 3'
(389..343); 99% identity.--DSS similarity to AA127694
zk89c11.rl Soares pregnant uterus NBHPU Homo sapiens cDNA
clone 490004 5' (77..125); 90% identity.--(15735..15713)
DSS similarity to W46603 zc32h10.rl Soares senescent
fibroblasts NBHSF Homo sapiens cDNA clone 324067 5'
(351..329); 100% identity.--(15735..15713) DDS similarity
to W46604 zc32h10.s1 Soares senescent fibroblasts NBHSF
Homo sapiens cDNA clone 324067 3' (299..321); 100%
identity."
misc_feature complement(15895..16082)
/note="predicted exon, program: grill2exons_human_1.3,
frame: 0, quality: excellent, score: 81.000--Other
overlapping matches:--(15936..15895) DDS similarity to
AA121532 zk89c11.s1 Soares pregnant uterus NBHPU Homo
sapiens cDNA clone 490004 3' (431..390); 100%
identity.--(15970..15895) DDS similarity to AA127694
zk89c11.rl Soares pregnant uterus NBHPU Homo sapiens cDNA
clone 490004 5' (1..76); 100% identity."
repeat_region 17926..18190
/rpt_family="Alu"
misc_feature complement(18270..18438)
/note="predicted exon, program: grill2exons_human_1.3,
frame: 0, quality: excellent, score:
100.000--(18438..18284) DDS similarity to W66776
mel17b11.rl Soares mouse embryo NBME13.5 14.5 Mus musculus
cDNA clone 387741 5' similar to PIR:B38252 B38252
granulocyte colony-stimulating factor receptor precursor
(157..1); 82% identity.--(18438..18406) DDS similarity to
AA049280 mj45d02.rl Soares mouse embryo NBME13.5 14.5 Mus
musculus cDNA clone 479043 5' similar to SW:IL6B_MOUSE
Q00560 INTERLEUKIN-6 RECEPTOR BETA PRECURSOR (432..464);
88% identity."
misc_feature complement(18540..18697)
/note="predicted exon, program: grill2exons_human_1.3,
frame: 1, quality: excellent, score: 100.000--DSS
similarity to W66776 mel17b11.rl Soares mouse embryo
NBME13.5 14.5 Mus musculus cDNA clone 387741 5' similar to
PIR:B38252 B38252 granulocyte colony-stimulating factor
receptor precursor (316..158); 92% identity.--DSS
similarity to AA049280 mj45d02.rl Soares mouse embryo

```



